Autonomous Road Driving Arenas for Performance Evaluation

Chris Scrapper, Stephen Balakirsky, Brian Weiss
National Institute of Standards and Technology
Intelligent Systems Division
Gaithersburg, Maryland 20899
Email: {scrapper, balakirsky, bweiss}@nist.gov

ABSTRACT— The development of performance metrics is critical in the evaluation and advancement of intelligent systems. Obtaining the pinnacle of intelligence in autonomous vehicles requires evolutionary standards and community support. In order to analyze and compare competing implementations of intelligent systems, the critical components of the system must be decoupled to facilitate repeatable trials that target specific aspects of the system's overall task. This paper presents a framework for a real virtual simulation environment that provides the facilities and tools to formally test the limitations and capabilities of autonomous road driving vehicles.

Keywords:
Performance Metrics, Autonomous Road Driving, Real Virtual Simulation, MOAST, RCS

> "A major barrier to the development of intelligent systems is the lack of metrics and quantifiable measures of performance. There cannot be a science of intelligent systems without standard units of measure." [1]

1. Introduction

Autonomous vehicles have made remarkable advancements during the past few decades. Researchers have produced many systems in the quest to develop intelligent control architectures capable of various autonomy levels. These advancements have elevated the capabilities of autonomous driving and have been showcased in many competitions and events recently. One instance of such competitions is the Defense Advanced Research Projects Agency's (DARPA) Grand Challenge [2], which seeks to advance a variety of architectures and implementations by luring researchers to demonstrate their autonomous systems in a timed, autonomous ground vehicle "race". However, a task's parameters and constraints may dictate that high speed is not always indicative of successful performance by an autonomous vehicle (e.g., military operations may encourage stealth over speed). Thus, autonomous systems should produce a multitude of behaviors within the same dynamic environment based upon a particular set of constraints. The success of the architecture relies heavily on the ability of the intelligent system to be cognizant of its environment. To accomplish this, intelligent systems must use perception systems to gather situational information from non-deterministic, stochastic real world environments. Following detection and classification of the environment, the autonomous system selects an appropriate mode of operation. On-road driving is the mode of operation that is targeted in this paper.

As the pace of this technology continues to develop and even accelerates, performance metrics must be designed and refined to measure the effectiveness of autonomous systems. An archetypical system will emerge from the application of these metrics as the eventual pinnacle of autonomous road driving achievement. By this definition, an autonomous system will have the ability to produce rational behaviors allowing it to successfully achieve its goals in any dynamic, unstructured, and complex environment. In order to realize this archetypical system, critical architecture components must be de-coupled for independent evaluation. The evaluation of these components will allow evolutionary standards to be applied as necessary to help define the performance metrics, direct the focus of the community, and encourage collaborative efforts. Without community-acknowledged performance metrics and a complete comprehension of capabilities and limitations, it will be very difficult to assess and compare competing methods [3].

This paper describes the testing, development, and evaluation of autonomous vehicles. In Section 2, there is a brief discussion on elements of autonomous road driving and the critical components of the intelligent system that governs the vehicles behavior. In Section 3, concept for the testing and evaluation of autonomous road driving systems is presented. Section 4 briefly summarizes the technical rationale behind this effort.

2. Autonomous Road Driving

As mentioned above, the archetypical autonomous vehicle is capable of operating within its domain for extended periods of time without human intervention, acting in a manner that escalates the likelihood of successful completion of the system's intent or goals [4]. However, as [5] expresses, designing an outdoor mobile robot to follow a straight line is one thing; operating in a realistic environment is another.

Real outdoor environments are stochastic, non-deterministic domains that contain countably infinite combinations of percepts or sequences of situations [6]. Even road networks, which provide some structure in the form of lane markings, signs, and rules that govern the road, are still extremely complex [7]. The road networks can be seen as multi-agent heterogeneous environments that contain both static and dynamic obstacles.

The rationality of autonomous vehicles within any domain depends on the performance measures delineating the criteria for success, *a priori* knowledge of the domain, plausible actions that can be performed, and *in situ* knowledge perceived from the environment [6], [8]–[10]. This requires that an autonomous road driving vehicle be a goal-directed agent that is capable of applying situational awareness to these multi-agent environments. For example, the vehicle must be able to navigate intersections, avoid obstacles, and follow the rules of the road in any weather or environmental state. This highlights the need for an autonomous system to contain a general theoretical model of intelligence that permits the integration of knowledge, perception, and behavior generation into a unified framework [11].

Albus *et al.* [12, Chap. 2] provides a comprehensive overview of reference architectures for intelligent systems. Yet, the realization of an autonomous vehicle operating safely and efficiently in a realistic on road driving domain has not occurred. The ability for the mobile robot to "act rationally" [6] in dynamic non-deterministic environments is a central aspect to the evaluation of autonomous road vehicles. [13] attributes the formulation of behaviors or actions to the convergence of the ontological models, or knowledge, with sequences of percepts. Thus, posing the ability to gather knowledge in dynamic environments that is accurate, reliable and current is necessary for the development of rational behaviors [12], [14], [15]. This has led to the identification of several key subsystems of an autonomous road driving system that will serve as the initial test elements to foster evolutionary standards and performance metrics to achieve the archetypical autonomous system. The components of the autonomous system that will initially be targeted are high-level control, low-level control, and perception.

High-level control is an aspect of an autonomous vehicle that provides the system with long-term goals given constraints on the system. Constraints are parameters that are incorporated while generating behaviors that dictate the selection of the appropriate routes and maneuvers on the road, e.g., find the shortest path while obeying the rules of the road. Therefore, high-level control is comprised of a set of functions that use artifacts derived from a knowledge base concerning the rules of the road and the degree of aggressivity to formulate coarse path plans and elementary maneuvers that the vehicle must perform to achieve its ultimate goal. When an obstruction or obstacle is encountered that renders the existing path plan obsolete, the high-level control system must be able to re-evaluate the situation to replan alternative routes and/or maneuvers.

Low-level control is an aspect of an autonomous vehicle that provides the system with more immediate actions or short-term goals. The low-level control system uses a coarse path plan that is received from the high-level control system to generate trajectories. The trajectories provide the system with fine tuned aspects of autonomy such as lane control and obstacle avoidance. In short, low-level control provides the actuators of the vehicle with kinetically viable trajectories that do not endanger the vehicle or other elements within the environment [16]. Therefore, low-level control must account for the presence of both static and dynamic obstacles and/or changing conditions in the environment when generating these trajectories.

Outdoor environments in the real world present an enormous perception problem, given varying weather conditions, shadows, lighting conditions, etc. [5]. Perception systems give the autonomous system the ability to detect the edge of roads, read traffic signs, and track objects within the environment. Sensory information is processed by the perception systems in order to create a dynamic internal representation of the real world environment. This representation is combined with ontological models to extract knowledge to be used in the generation of appropriate actions. This extracted knowledge is then combined with *a priori* knowledge of the mode of operation to identify, recognize, and predict objects within a time horizon of the vehicle [12].

3. Program Concepts

The Intelligent System Division (ISD) at the National Institute of Standards and Technology (NIST) has long been a proponent of performance metrics for intelligent systems. Under the DARPA Mobile Autonomous Robot Software (MARS) program [17], [18], ISD is proposing to construct a set of Autonomous Road Driving Arenas (ARDA) to encourage the development of autonomous road driving vehicles and the creation of community recognized performance metrics for autonomous road driving. The sharing of knowledge by the community has facilitated rapid advancement of technologies and performance metrics within its domain. Therefore, this effort will attempt to gather the participation of academic, government, and commercial institutions through publications, technical reports, data sets, competitions, and accessibility to the arenas themselves.

This effort will draw on experience from the development and proliferation of the NIST Reference Test Arenas for Autonomous Mobile Robots [19] and the associated formal testing of urban search and rescue robots. The Urban Search And Rescue (USAR) arenas originally provided physical arenas to conduct controlled repeatable scenarios that challenge the different facets of the USAR robots. Since the inception of the USAR arenas, virtual components and tools have been developed to simulate these arenas [20]. Under the Mobility Open Architecture Simulation and Tools (MOAST) [21], efforts are being made to seamlessly integrate the real and virtual simulation environments into a single simulation environment and toolset. In recent years, the arenas have proven to be an invaluable resource, lifting the standards of development within the community by providing comprehensive data sets, publicly accessible USAR arenas, and competitions such as RoboCup Rescue.

The Autonomous Road Driving Arenas (ARDA) will provide an environment that will support the cost effective development and testing of critical components of autonomous road driving vehicles. The ability to measure performance in terms of efficiency, effectiveness, precision, and reliability requires that the testing environment have the ability to independently evaluate the various aspects of autonomous road driving through various methods as detailed by [22]. MOAST provides the reference framework for the development of the facilities and tools to gather vital information during repeatable trials, which can be used to measure the performance of the autonomous vehicles both quantitatively and qualitatively.

ARDA will provide a realistic, scaled world that will evolve from simple lanes to a controlled dynamic environment that will exploit the capabilities and limitations of autonomous road driving vehicles. Therefore, a standardized platform will also be developed that will support these testing efforts by providing researchers with identical physical capabilities for their algorithms to perform against. It is important that ARDA provides the assurance that all participants experience the same degree of difficulty in order to measure and compare the performance of intelligent autonomous vehicles [1].

The capabilities of ARDA will emerge in three phases. Phase I will consist of a 2D planar arena with simple controlled intersections and scripted traffic. In Phase II, more complex environmental conditions, such as dynamic changing lighting conditions, and more challenging traffic conditions, such as construction zones, will be integrated into ARDA. Phase III will complete ARDA by adding non-planar road surfaces and other obscuring features such as bridges, hills, and tunnels. At the end of each phase, a technical report will be issued detailing the developmental life cycle of each particular phase, including feasibility analysis and technology assessment studies.

4. Technical Rationale

The rich world that exists within virtual simulation environments has long been an invaluable resource for the testing and development of intelligent systems. Traditionally, these environments were developed in-house for specific purposes [23]. More recently, there has been an emergence of game-based simulation environments that capitalize on the optimized rendering facilities and rich dynamic physic models contained in professionally developed game engines [24]. The Serious Games Initiative, USC's Institute for Creative Technologies, the MOVES Institute sponsored by the US military and many others, illustrate the rapid growth of the interest in the gaming community for robust simulation packages.

Though the rich virtual worlds contained within these simulation environments are invaluable, they are still not the "magic bullet". Algorithms that have been trained or developed within the virtual environments can become dependent on synthetic data cues. The majority of the virtual simulation environments are built on deterministic models that do not provide false alarms or missed detection and employ unrealistic mobility models. However, virtual environments do provide the flexibility and control in the environment to evaluate the capabilities and limitations of autonomous systems. This facilitates the ability to target specific aspects of autonomous systems by decoupling critical components and logging vital data of an autonomous system in repeatable trials.

Russell *et al.* [6] describes the environments contained in the real world as stochastic, non-deterministic, dynamic, multi-agent environments that are only partially observable to any one agent. When implementing autonomous systems on the actual platforms, the system must contend with real mobility characteristics of the vehicle's platforms. The system must also rely on real sensors that provide noisy data and missed detection. Thorpe *et al.* [5] express that these environments force the development of algorithms that are more robust and reliable in order for the system to contend with the realistic mobility characteristics of the actual platform and imperfect sensory data. However, the real world environments also have drawbacks. Since the planner relies heavily on the information obtained by the perception system, the performance of the system is greatly limited by what is known as the myopic planning effect [25]. In short, the myopic planning effect is the limitation of the planner to anticipate objects beyond the range of the vehicle mobility sensors, e.g. LADAR (laser radar). The inherent stochastic nature of real world environments makes it difficult to perform repeatable trials that have the same environmental state and conditions. Lastly, there is a great concern to adhering to Isaac Asimov's "Three Laws of Robotics" (safety to itself and others) when using the real world environments for the testing and evaluation of autonomous vehicles.

The technical rationale behind this effort is to create an environment that benefits from the best of both real and virtual worlds and to create an array of incremental steps that provide a smooth gradient to transition an autonomous system from a purely virtual world to a entirely real world. By allowing the virtual components and the real components to function where they perform best, ARDA seeks to achieve an optimal testing environment that seamlessly integrates both real and virtual components [21], [26], [27]. For example, Barbera *et al.* [28] describes a methodology of reverse engineering performance metrics for sensors and perception systems by evaluating the world model and knowledge requirements needed for the behavior generation systems. Therefore, in the ARDA environment the use of virtual sensing can give us the opportunity to understand what is better needed from a perception system. ARDA will use the MOAST general-purpose framework to provide this effort with a developmental reference model during implementation.

MOAST manages the fidelity of the simulated worlds in order to reduce the computational overhead that is inherent in dynamic simulation environments. This is accomplished by the concurrent use of the real vehicle platform, a high-fidelity simulator and a low-fidelity simulator. An example of the implementation of the MOAST framework is to use the vehicle platform to run the low-level mobility, planning and low-level perception, using a high fidelity simulator to accurately simulate objects in the vehicles immediate sphere of influence, and using a low-level simulator for tracking, and simulating distant objects. Objects are automatically transferred between simulators as their proximity to the vehicles changes within the MOAST framework. This framework also permits the transparent transference of data between real and virtual components, which gives the developer the ability to toggle a "switch" between the two components. For example, virtual sensor that simulate raw sensory data or real sensors can be used as input into a perception system which in turn can either be an actual functioning subsystem or a virtual system that simple places preprocessed knowledge into the *WM.*

ARDA will provide an environment, with extensive high-resolution data sets of annotated maps, features, and *a priori* knowledge to implement the MOAST framework. This will provide ground truth and a benchmark for the testing, development, and evaluation of the intelligent systems in this specific domain. A scaled physical model of a road network will be used to initially evaluate high-level planning, low-level mobility, and low-level perception. Later, this environment will be used in the testing and development of multiple agent systems (MAS) [24], essential components such as various levels of Situational Awareness [29], and axioms of cooperation such as distributed artificial intelligence, resource conflicts, learning, group architectures, and geometric problems [30] .

ARDA will leverage existing technologies and on-going efforts in simulation to construct a simulation environment based on the MOAST framework consisting of four basic components: Real World Arenas, Virtual Simulation Worlds, Vehicle Platforms, and a Tracking System. The real world arenas will be the physical world where the vehicle platforms will be used to test various aspects of autonomous systems. The virtual simulation environment will reflect the real world environment and will utilize information captured by the tracking system in order to provide ARDA with an array of visualization tools and services. Tracking will also be used to provide navigation and other vital data to the agents operating within the arenas. These four components will be detailed further in the following subsections.

4.1. *Real World Arenas*

ARDA implements the MOAST framework in a one-tenth scaled replica of a real environment that includes objects and features such as plants, buildings and roads. The arena design will provide an equal level of difficulty for every autonomous system operating in ARDA. The arenas will target specific aspects of the autonomous system by mimicking the environmental conditions found in the on-road driving domain. These environmental conditions include static and dynamic objects, variable weather and lighting conditions, and features associated with road networks, e.g. controlled intersections and lane markings. The initial configuration of the arena will consist of a planar world with limited environmental conditions. More complex environmental conditions will be incrementally introduced in response to the escalating capabilities of the autonomous road driving systems.

Essential design factors that will continually be considered will be the use of commercially available components, modularity, reconfigurability, and durability. Commercially available components are standardized parts that are readily available and relatively inexpensive. The use of these components will assist in the proliferation of the arenas by allowing participants to construct or repair their own replicas according to ARDA's specs. The consistency of the standardized components expedites rapid repairs while maintaining the arena's integrity. A modular design also assists in efficient repairs through the easy replacement of damaged components. Reconfigurability is the ease with which environmental features can be added, modified, or deleted from the arenas. This gives ARDA the versatility to alter the configuration of its constituent components in order to target different aspects of the autonomy. Reconfigurability facilitates different road network implementations ranging from small demonstration configurations to large competition networks. Durability refers to the resilience of the arenas to the continual weathering due to the environmental conditions and the stress of assembly/disassembly and packing/unpacking associated with frequent travel.

With the design criteria in mind, ARDA will be composed of inter-locking rectangular frames with attached panels to create the base, upon which the road networks will be placed. The primary members of the road network frames will potentially be aluminum extrusion. Several types of commercially available aluminum extrusion are easily machined to provide unlimited attachment solutions. Plastic panels that are durable and lightweight will be tailored to conform to the frame and will provide the road surface and the foundation upon which terrain features will be placed. Upon the construction of the base, various terrain objects and artificial turf will be used to populate the world. Lane markings will be painted on the road surface to create passing and non-passing zones. Controlled intersections will be composed of traffic signs and/or traffic lights. Traffic lights will be controlled through basic transistor logic or through interfaces to the virtual simulation controller.

4.2. *Vehicle Platforms*

All vehicle platforms that conform to the one-tenth scaled environment will be permitted to compete in ARDA. However, standardized platforms will be developed for use in ARDA that are akin to modified remote control cars (popular with hobbyists). For these platforms to behave rationally, the platforms must extend their computational capabilities to achieve the capabilities and computational specs currently available in an implementation of an actual autonomous vehicle, such as the NIST High Mobility Multipurpose Wheeled Vehicle (HMMWV) [31]. This requires that the platforms contain sufficient computational power, control architecture, sensors, and communication bandwidth. Initially, a few criteria will be used to evaluate the platform's design. The criteria identified to maximize the platform's capabilities are:

- Reference Control Architecture
- Standardized Parts
- Computational Power
- Battery-Life vs. Power Consumption
- Communication
- Sensors
- Closed-Loop Controller for Actuators, Sensors, and Communication

1) Reference Control Architecture: It is critical to identify a reference model architecture to organize the hardware and software within an intelligent system. The Real-Time Control System architecture, RCS, was chosen for the platforms in the ARDA because it uses a modular, hierarchical control architecture as an efficient way to manage the system's complexity. RCS is a real-time, distributed, hierarchical architecture that consists of computational nodes with well-defined functional components and clearly defined interfaces. RCS contains a systematic regularity so that each control node in the hierarchy performs the same general type of functions: sensory processing (*SP*), world modeling (*WM*), value judgment (*VJ*), and behavior generation (*BG*).

The principal difference between control nodes at the same level is in the set of resources managed, while the principal difference between nodes at different levels is in the knowledge requirements and the fidelity of the planning space. This regularity in the structure enables flexibility in the system architecture that allows scaling of the system to any arbitrary size or level of complexity. [1]

Each level in RCS has a characteristic range and resolution in space and time, determined by the specific implementation requirements. Each level has characteristic tasks and plans, knowledge requirements, values, and rules for decision-making. Every module in each level has a limited span of control, a limited number of tasks to perform, a limited number of resources to manage, a limited number of skills to master, a limited planning horizon, and a limited amount of detail with which to cope [32].

2) Standardized Parts: The standardized vehicle platforms developed for ARDA will be modified one-tenth scaled model remote control (*RC*) cars. Standard off-the-shelf parts, e.g. chassis, wheels, suspension and servos/actuators, will provide the base hardware for the vehicle platforms. These parts for these RC cars are widely available, inexpensive, and standardized allowing for vehicles to be modified easily. This facilitates efficient repairs and construction of vehicles platforms, which is ideal for development and testing of this nature. This effort will capitalize on available internal programs, competitions, and partnerships to develop a standardized platform that will provide the optimal functionality within ARDA.

3) Computational Power: Each platform must have sufficient computational power to handle the computational load of the autonomous systems, e.g. low-level planner, high-level planner, and low-level perception systems. The MOAST framework and modular hierarchical design of RCS allows for standard low-power single board core-processing units to be coupled through the designated interfaces with off-board processing power to meet any computational load needed by the vehicle. This provides the environment the flexibility to distribute the autonomous system's components or subsystems between the onboard and off-board processing units.

4) Battery-life vs. Power Consumption: The vehicle platforms will not be tethered; therefore, their electrical systems must operate on battery power. This bounds the time that the vehicles will be able to operate in ARDA. This requires that battery packs provide a significant number of milli-amp hours (mAH) to effectively operate the vehicle's electrical system, without signifigant power loss, for a duration of time that is appropriate for testing. For example, the core-processing units available are 12 W systems, drawing 1.75 V. RC battery packs are available in a 6-cell battery back (9 V) that provide 3000 mAh, which equate to 3 Ah. Given the power requirements of the core-processing unit and additional TTL logic required, the system would draw one amp from the battery pack. This only gives the vehicle three hours of operation before the power supply would need to be recharged. This does not take into account the power requirements to operate the motor, servos, or additional hardware that may need to be added for sensing. Multiple batteries can be mounted on the platform that provides ample power for the vehicle's electrical system.

5) Communication: Again, the combination of the MOAST framework and RCS lends flexibility to distribute the components of the autonomous system between the onboard and off-board processing units. The interfaces between the distributed subsystems, command and control channels, require enough communication bandwidth to accommodate a wide range of command messages. The communication infrastructure of ARDA must also be able to handle the command and control interfaces for multiple vehicles and broadcast global information, such as navigation data.

6) Sensors: Intelligent systems gather knowledge about the state of the world and its relation to the world through sequential percepts captured through sensors. Sensing capabilities within the MOAST framework can either be obtained virtually or through real onboard sensors. These capabilities include the collection of navigation data, color video, and range finding information. In the scaled environment of ARDA, objects will remain proportionally scaled to their real world counterparts, requiring sub-centimeter resolution in the sensory information. During the first phase of the development, the vehicle platforms will rely heavily on virtual sensing. The vehicle platforms will be designed with appropriate interfaces that will facilitate the addition and/or re-configuration of real onboard sensors.

7) Closed-Loop Controller for Actuators, Sensors, and Communication: The architecture controlling the sensors, actuators and communications will rely on a core-processing unit for its computational requirements. Closed-loop adaptive controllers will provide the core-processing unit with control interfaces and added computational power to regulate its facilities. This provides each distinct facility with a closed communication control loop that is comprised of a command sent to the controller, feedback received from the controller, and data transference between the core-processing unit and the controlled facility. Each control loop will consist of a stacked I/O card that contains additional computational power to run small control programs.

4.3. Virtual Simulation Envirnoment

Within the MOAST architecture, virtual and real components are combined transparently to mutually augment each other in order to achieve an environment conducive for the evaluation of autonomous road driving vehicles. This effort requires the virtual simulation environment to be capable of handling the overhead of simulating complex, unstructured environments contained in road networks. The virtual environments will be required to provide multiple levels of resolution, e.g., in the sphere of influence around the vehicle the virtual simulator must have high fidelity, while the distant objects require less fidelity.

The virtual simulation environments must provide the facilities and interfaces to allow for the developer to create new worlds, provide virtual sensing capabilities, modify objects, and appropriate command and control channels to interface with all agents in both the real and virtual environments [33], [34]. The virtual world must also provide the facilities to log vital information in real-time and visualization tools to access and display the internal state of the vehicle during operation. These tools provide developers with a level of control that permits for the internal state of the vehicle to be modified. For example, a vehicle will perceive virtual objects as actual objects in the real world, where the virtual world will view the vehicle as another agent within the virtual world. Having these tools to visualize the internal state of the autonomous system in real-time gives the developer more insight into the decision making process and the causality of behaviors being produced.

The virtual simulation environments must also be able to provide realistic mobility characteristics, sensory data, and tactical behaviors for each vehicle. Dixon *et al.* [26] points out that a signifigent amount of low-level infrastructure is required to support the communication, control, and simulation of multiple agents in high-fidelity virtual environments. The low-level infrastructure is time-consuming and requires that the developer have a high-level of expertise to maintain the infrastructure [35]. It appears that there is a positive, non-linear correlation between the number of objects in the world and the computational load to simulate these objects. An effective way to manage the computational complexities within these virtual environments is to couple a high-fidelity simulator with a low-fidelity simulator. The high-fidelity simulation provides very accurate information about objects in a bounded region, where as the low-fidelity simulation simulates the behavior and sensing of a larger area of distant objects at a lower resolution. This provides the autonomous system with *in situ* knowledge at an appropriate resolution, while orchestrating the entire scenario at a lower, more efficient resolution.

4.4. Tracking System

A tracking system will be integrated in the arenas to collect real-time navigation data of all the subjects in the real world environment and broadcast this knowledge to all the constituents that comprise ARDA. This knowledge can serve as a virtual positioning sensor for each vehicle or can be coupled with *a priori* ground truth knowledge to provide other virtual sensing capabilities. It can provide redundant navigation information for each platform by being coupled with wheel encoders or other positioning sensors located on the actual vehicle. The tracking system logs this data so that is can be used for the evaluation of the behaviors generated by the autonomous vehicles. This positioning knowledge can be compared with positioning information logged from the vehicle to validate trajectories and coarse path plans. These options are being explored based upon the available platform sensors and current resolution, accuracy, and update rate limitations of available tracking systems.

Overhead, two-dimensional camera tracking is being considered as the initial tracking system due to the planar configuration in ARDA's real world environment. The technologies available in commercially available color video cameras/camcorders provide sufficient resolution and updates rates to serve as the overhead visual sensor for the tracking system. Many publicly available vision algorithms can be mated with most overhead cameras to provide tracking update rates of 30 Hz or slightly greater. One such available software package is Mezzanine [36]. This highly competent overhead visual tracking software is publicly available at no cost and may be used with almost any video camera/camcorder. The software enables the camera to track objects by following color-coded fiducials placed on top of each autonomous vehicle platform. Mezzanine provides real-time translation of pixel location into world coordinates.

Other software is widely available for camera tracking including algorithms developed by the RoboCup Soccer Small-Size Teams for use in their annual competitions [37]. This league features teams of five robots playing soccer against one another where all necessary sensing and processing takes place off-board. Each team sets up an overhead camera and places their own color-coded fiducials on their robots. Teams utilize their own tracking algorithms so that they may locate the position and orientation of their robots, their opponents, along with the position of the ball. The off-board computational facilities process this data and send the appropriate commands back to their robots. Several teams have extensively documented their vision algorithms that have the potential to be used with ARDA [38]–[40].

5. Summary

Under the DARPA MARS project, the Intelligent Systems Division of the National Institute of Standards and Technology is proposing the construction of a real/virtual simulation environment for the performance evaluation of autonomous road driving vehicles, ARDA. The effort is seeking to coalesce together an autonomous road driving community that will support the development, testing, and evaluation of various implementations of autonomous systems. ARDA will be designed to support the MOAST framework, which will give the developers the ability to target specific aspects of autonomous road driving in repeatable scenarios. This effort hopes for the realization of an intelligent autonomous system that will service as the archetypal system, which all other systems will be modeled and compared.

6. Acknowledgement

This work was supported in part by the Army Research Lab's program in unmanned vehicles (PM. C. Shoemaker) and DARPA's Mobile Autonomous Robotic Software program (PM. D. Gage).

7. Product Disclaimer

Commercial equipment and materials are identified in order to adequately specify certain procedures. In no case does such identification imply recommendation or endorsement by the National Institute of Standards and Technology, nor does it imply that the material or equipment identified are necessarily the best available for the purpose.

8. REFERENCES

[1] J. Albus, "Metrics and Performance Measures for Intelligent Unmanned Ground Vehicles," in *Proceedings of the 2002 Performance Metrics for Intelligent Systems Workshop* , 2002.

[2] DARPA, "Grand Challenge," 2004. [Online]. Available: http://www.darpa.mil/grandchallenge

[3] J. Evans and E. Messina, "Performances Metrics for Intelligent Systems," in *Proceedings of the 2000 Performance Metrics for Intelligent Systems Workshop* , 2000.

[4] H.-M. Huang, "Terminology for Specifying the Autonomy LEvels for Unmanned Systems: Version 1.0," NIST, NISTIR 1011, 2004.

[5] C. Thorpe, M. Hebert, T. Kanade, and S. Shafer, "Toward Autonomous Driving: the CMU Navlab. I - Perception," *Expert, IEEE [see also IEEE Intelligent Systems]*, 1991.

[6] S. Russell and P. Norvig, *Artificial Inelligence: A Modern Approach*. Prentice-Hall, 1995.

[7] D. Pomerleau and T. Jochem, "Image Processor Drives Across America," *Photonics Spectra*, 1996.

[8] S. Balakirsky and A. Lacaze, "World Modeling and Behavior Generation for Autonomous Ground Vehicles," in *Proceedings of the IEEE Conference on Robotics and Automation*, 2000.

[9] S. Balakirsky and A. Lacaze, "Value-Driven Behavior Generation for an Autonomous Mobile Ground Robot," in *Proceedings of the SPIE 16th Annual Symposium on Aerospace—Defense Sensing, Simulation and Controls*, 2002.

[10] J. Albus, "Outline for a Theory of Intelligence," *IEEE Transaction on Systems Man and Cybernetics*, 1991.

[11] A. Meystel and J. Albus, *Intelligent Systems: Architecture, Design, and Control*, ser. Wiley Series on Intelligent Systems. John Wiley and Sons, Inc., 2002.

[12] J. Albus and A. Meystel, *Engineering of Mind: An Introduction to the Science of Intelligent Systems*, J. Albus and A. Meystel and L. Zadeh, Ed. John Wiley and Sons, Inc 2001, 2001.

[13] N. Howard, *The Theory of Intention Awareness in Tactical Military Intelligence: Reducing Uncertainty by Understanding*. First Book Publisher, 2002.

[14] S. Balakirsky, *A Framework for Planning with Incrementally Created Graphs in Attributed Problem Space*. IOS Press, 2003.

[15] B. Krogh and C. Thorpe, "Integrated Path Planning and Dynamic Steering Control for Autonomous Vehicles," in *Proceedings of 1986 IEEE International Conference on Robotics and Automation (ICRA '86)*, April 1986, pp. 1664 – 1669.

[16] C. Thorpe, M. Hebert, T. Kanade, and S. Shafer, "Toward Autonomous Driving: the CMU Navlab. II - Architecture and Systems," *Expert, IEEE [see also IEEE Intelligent Systems]*, 1991.

[17] D. Gage, "Mobile Autonomous Robot Software (MARS)," 2003. [Online]. Available: http://www.darpa.mil/ipto/programs/mars/index.htm

[18] NIST, "DARPA Mobile Autonomous Robot Software," 2003. [Online]. Available: http://www.isd.mel.nist.gov/projects/darpa_mars

[19] A. Jacoff, B. Weiss, and E. Messina, "Evolution of a Performance Metric for Urban Search and Rescue Robots," in *Proceedings of the 2003 Performance Metrics for Intelligent Systems Workshop*, 2003.

[20] J. Wang, M. Lewis, and J. Gennari, "USAR: A Game Based Simulation for Teleoperation," in *Proceedings of the 47th Annual Meeting of the Human Factors and Ergonomics Society*, 2003.

[21] S. Balakirsky and E. Messina, "A Simulation Framework for Evaluating Mobile Robots," in *Proceedings of the 2002 Performance Metrics for Intelligent Systems Workshop* , 2002.

[22] NIST, "DARPA Mobile Autonomous Robot Software," 2003. [Online]. Available: http://www.isd.mel.nist.gov/projects/darpa_mars/Perf_Metrics.html

[23] R. Sukthankar, D. Pomerleau, and C. Thorpe, "SHIVA: Simulated Highways for Intelligent Vehicle Algorithms," in *Proceedings of Intelligent Vehicles '95*, September 1995, pp. 332 – 337.

[24] R. Adobbati, A. Marshall, A. Scholer, S. Tejada, G. Kaminka, S. Schaffer, and C. Sollitto, "Gamebots: A 3D Virtual World Test-Bed For Multi-Agent Research ," in *Proceeding of the 2nd Workshop on Infrastructure for Agents, MAS, and Scalable MAS at Autonomous Agents*, 2001.

[25] B. Nabbe and M. Hebert, "Where and When to Look," in *IROS 2003*. IEEE, October 2003.

[26] K. Dixon, J. Dolan, W. Huang, C. Paredis, and P. Khosla, "RAVE: A Real and Virtual Environment for Multiple Mobile Robot Systems," in *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS'99)*, vol. 3, October 1999, pp. 1360–1367.

[27] R. Vaughan, B. Gerkey, and A. Howard, "On Device Abstractions for Portable, Reusable Robot Code," in *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems*, ser. IROS 2003, 2003.

[28] A. Barbera, J. Horst, C. Schlenoff, E. Wallace, and D. Aha, "Developing World Model Data Specifications as Metrics for Sensory Processing for On-Road Tasks," in *Proceedings of the 2003 Performance Metrics for Intelligent Systems Workshop* , 2003.

[29] M. Endsley and R. R. Hoffman, "The Sacagawea Principle," *Intelligent Systems, IEEE*, 2002.

[30] Y. Cao, A. Fukunaga, and A. Kahng, "Cooperative Mobile Robotics: Antecedents and Direction," *Autonomous Robots*, vol. 4, pp. 7–27, 1997.

[31] K. Murphy, M. Abrams, S. Balakirsky, D. Coombs, T. Hong, and S. Legowick, "Intelligent Control for Unmanned Vehicles," in *Proceeding of the 2000 World Automation Congress Conference*, 2000.

[32] J. Albus, H. Huang, E. Messina, K. Murphy, M. Juberts, A. Lacaze, S. Balakirsky, M. Shneier, T. Hong, H. Scott, F. Proctor, W. Shackleford, J. Michaloski, A. Wavering, T. Kramer, N. Dagalakis, W. Rippey, K. Stouffer, S. Legowik, R. Bostleman, R. Norcross, A. Jacoff, S. Szabo, J. Falco, B. Bunch, J. Gilsinn, T. Chang, A. Meystel, A. Barbera, M. Fitzgerald, M. DelGiorno, and R. Finkelstein, "4D/RCS Version 2.0: A Reference Model Architecture for Unmanned Vehicle Systems," NIST, NISTIR 6910, 2002.

[33] B. P. Gerkey, R. T. Vaughan, and A. Howard, "The Player/Stage Project: Tools for Multi-Robot and Distributed Sensor Systems," in *Proceedings of the International Conference on Advanced Robotics*, ser. ICRA 2003, 2003, pp. 317–323.

[34] G. A. Kaminka, M. M. Veloso, S. Schaffer, C. Sollitto, R. Adobbati, A. N. Marshall, A. Scholer, and S. Tejada, "Gamebots: A Flexible Test Bed for Multiagent Team Research," *Communications of the ACM*, vol. 45, no. 1, pp. 43–45, January 2002.

[35] A. Jacoff, E. Messina, and J. Evans, "Performance Evaluation of Autonomous Mobile Robots," *Industrial Robot: An International Journal*, 2002.

[36] USC Robotics Resesarch Lab, "Mezzanine: An Overhead Visual Object Tracker," 2003. [Online]. Available: http://playerstage.sourceforge.net/mezzanine/mezzanine.html

[37] B. Browning, "Official RoboCup Small Size League," 2003. [Online]. Available: http://www-2.cs.cmu.edu/ brettb/robocup

[38] A. Egorova, A. Gloye, A. Liers, R. Rojas, M. Schreiber, M. Simon, O. Tenchio, and F. Wiesel, "FU-Fighters 2003 (Global Vision)," in *RoboCup-2003 - Proceedings of the International Symposium*, 2003.

[39] J. Bruce, T. Balch, and M. M. Veloso, "Fast and inexpensive color image segmentation for interactive robots," vol. 3, October 2000.

[40] J. Loomis, J. Palmer, and P. Pandit, "Performance Development of a Real-Time Vision System," Master's thesis, Cornell University, 2003.